Tensor decomposition of polarized seismic waves

Francesca Raimondi, Pierre Comon

Tensor decomposition of polarized seismic waves

Francesca Raimondi, Pierre Comon*

GIPSA-Lab, UMR5216,
11 rue des Mathematiques, Grenoble Campus, BP.46, F-38402 St Martin d'Heres cedex, France
firstname.lastname@gipsa-lab.grenoble-inp.fr

Résumé – En traitement d'antenne, les décompositions tensorielles permettent d'estimer conjointement les sources et de les localiser. Pour que ces dernières puissent être utilisées, il faut que les données présentent au moins trois diversités, qui sont habituellement le temps, l'espace, et la translation dans l'espace. L'approche présentée ici est basée sur la diversité de polarisation, une alternative très attractive lorsque l'antenne ne jouit pas d'invariance spatiale. Nous dérivons ensuite les bornes de Cramér-Rao dans ce contexte, en nous appuyant sur des conventions de différentiation de variables mixtes réelles et complexes.

Abstract – In antenna array processing, tensor decompositions allow to jointly estimate sources and their location. But these techniques can be used only if data are recorded as a function of at least three diversities, which are usually time, space and space translation. The approach presented therein is based on polarization diversity, a very attractive alternative when the antenna array does not enjoy space invariance. Then we derive Cramér-Rao bounds in this context, by resorting to differentiation conventions for real-complex mixed variables.

1 Introduction

Starting from the ideas on *vector sensor array* developed for seismic waves in [1] from the more general model for polarized waves in [2] and [3], we state the observation model in tensor form. Next we compute the Cramér-Rao Bounds (CRB) for the joint estimation of the four parameters of polarized seismic waves. The ultimate estimation performances are then compared to the CRB as a function of the Signal to Noise Ratio (SNR). A deterministic approach based on tensor decomposition has been introduced in [4]. The advantage of tensor decompositions lies in the need for shorter data records, since the estimation of statistical quantities from available samples is not a requirement (as opposed to traditional high resolution algorithms such as MUSIC [5] and ESPRIT [6]). CRB for the low-rank decompositions of multidimensional array was derived in [7] and extended in [8]. Polarization of waves has been first introduced in [9] as a multidimensional diversity in the tensor approach. The same authors in [10] explore the concept of polarization separation and its influence on performances.

Notation We shall assume throughout the following notations: matrices, column vectors and scalars are denoted respectively in bold uppercase, *e.g.* $\boldsymbol{A}$, bold lowercase, *e.g.* $\boldsymbol{v}$, and plain lowercase; in particular array entries are written *e.g.* v_j or A_{ij}. Transposition, complex conjugation and Hermitian transposition are denoted by $(^{\mathsf{T}})$, $(^*)$ and $(^{\mathsf{H}})$, respectively. Arrays with more than two indices are referred to as tensors, with some abuse of terminology [11], and are denoted in bold calligraphic, as $\mathcal{T}$. The outer (tensor) product between two vectors is denoted by $\boldsymbol{u}\otimes\boldsymbol{v}$. Finally, $\|\cdot\|_F$ refers to the Frobenius norm; $\boxtimes$ denote Kronecker product. For the sake of conciseness, $\boldsymbol{a}_i$ will represent the i-th column of matrix $\boldsymbol{A}$.

*This work has been supported by the ERC project "DECODA" no. 320594, in the frame of the European program FP7/2007-2013.

2 Physical model

2.1 A 4-parameter far-field model

The physical quantity measured is the *particle displacement vector* recorded by a three-component particle displacement sensor (or geophone), located at a given point in space, in the direction of the x-, y- and z-axes of its reference system. The z-axis is required to be perpendicular to the earth's surface. The following parameterization is based on the definition of four angular parameters. First, the unit vector pointing to the source is given by

$$\boldsymbol{u} = \begin{bmatrix} \cos\theta\ \cos\psi \\ \sin\theta\ \cos\psi \\ \sin\psi \end{bmatrix}$$

where $\theta \in (-\pi, \pi]$ refers to the azimuth and $\psi \in [-\pi/2, \pi/2]$ to the elevation of the source. Second, the polarization ellipse is described by the orientation angle $\alpha \in (-\pi/2, \pi/2]$ and the ellipticity angle $\beta \in [-\pi/4, \pi/4]$. Two models can be drawn respectively for *transverse* (TR)

waves and *tilted generalized Rayleigh* (TGR) waves [1]. For a transverse wave, the polarization ellipse lies in the space orthogonal to the direction of propagation $\boldsymbol{u}$, and is spanned by the columns $\{\boldsymbol{v}_1\, \boldsymbol{v}_2\}$ of the matrix:

$$\boldsymbol{V}^{TR} = [\boldsymbol{v}_1\,,\ \boldsymbol{v}_2] = \begin{bmatrix} -\sin\theta & -\cos\theta\ \sin\psi \\ \cos\theta & -\sin\theta\ \sin\psi \\ 0 & \cos\psi \end{bmatrix} \tag{1}$$

In the TGR model, the polarization ellipse is confined in the plane spanned by the columns of matrix:

$$\boldsymbol{V}^{TGR} = [\boldsymbol{u}\,,\ \boldsymbol{v}_2] = \begin{bmatrix} \cos\theta\ \cos\psi & -\cos\theta\ \sin\psi \\ \sin\theta\ \cos\psi & -\sin\theta\ \sin\psi \\ \sin\psi & \cos\psi \end{bmatrix}$$

If the complex envelope of the source signal is denoted by $s(t)$, the general data model can be written as

$$\boldsymbol{y}(t) = \boldsymbol{V}(\theta,\psi)\,\boldsymbol{Q}(\alpha)\,\boldsymbol{w}(\beta)\,s(t) \in \mathbb{C}^{3\times 1} \tag{2}$$

where $\boldsymbol{V}(\theta,\psi)$ is one of the above matrices,

$$\boldsymbol{Q} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}, \quad \boldsymbol{w} = \begin{bmatrix} \cos\beta \\ \imath\,\sin\beta \end{bmatrix}$$

in the absence of noise, and $\imath = \sqrt{-1}$.

2.2 Seismic waves and polarization

There exist several types of elastic waves associated with seismic activities [1]. *Primary* waves (or P-Waves) are compressional elastic waves whose particle displacement vector is parallel to the direction of propagation. For these waves, $\alpha = \beta = 0$, which leads to a linearly polarized wavefront with particle motion along the direction of propagation $\boldsymbol{u}$. *Rayleigh Waves* are elliptically polarized *surface* waves. Therefore, $\psi = 0$, provided that the xy-plane corresponds to the earth surface. For these waves, it is obvious that $\alpha = 0$ and then $\boldsymbol{Q} = \boldsymbol{I}$. *Secondary* or *Shear* waves (or S-waves) are transverse elliptically polarized in general. P-Waves and Rayleigh Waves are particular cases of the TGR model described in Section 2.1 for elliptically polarized waveforms: the direction of propagation is located in the plane spanned by the ellipse major and minor axes. For reasons of space, we shall concentrate on *TR waves* in the remainder.

2.3 Tensor model

Now suppose that data are recorded on K polarized sensors located at points in space defined by vectors $\boldsymbol{g}(k) := [g_k^x; g_k^y; g_k^z] \in \mathbb{R}^3$, $1 \le k \le K$. Also suppose that R far-field narrow-band sources impinge on this vector sensor array from direction $\boldsymbol{u}(r)$, $1 \le r \le R$, and denote ω their common angular pulsation. We make the assumption that impinging waves have elliptical polarization (neither linear nor circular). Then from (2) we can assume the following observation model in baseband about pulsation ω:

$$\mathcal{T} = \mathcal{Z} + \mathcal{E}, \quad \mathcal{Z} = \sum_{r=1}^{R} \boldsymbol{a}(r) \otimes \boldsymbol{b}(r) \otimes \boldsymbol{s}(r) \tag{3}$$

where $a_k(r) = \frac{1}{K}\exp\left\{\imath\frac{\omega}{v}\boldsymbol{g}(k)^{\mathsf{T}}\boldsymbol{u}(r)\right\}$ is the k-th entry of the steering vector, v the wave celerity, $\boldsymbol{b}(r) = \boldsymbol{V}(\theta_r,\psi_r)\boldsymbol{Q}(\alpha_r)\boldsymbol{w}(\beta_r) \in \mathbb{C}^{L\times 1}$ $(L = 3)$ characterizes the propagation type, (θ_r,ψ_r) refers to the Direction of Arrival (DoA) of the r-th source and (α_r,β_r) its polarization, and $s_m(r)$ is the signal propagating from the r-th source and received at time t_m, $1 \le m \le M$. The additive noise $\mathcal{E}$ is assumed to be i.i.d. circular Gaussian and independent of the sources. In terms of arrays of coordinates, model (3) rewrites:

$$Z_{k\ell m}(\boldsymbol{\theta},\boldsymbol{\psi},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{S}) = \sum_{r=1}^{R} a_k(r)b_\ell(r)s_m(r) \tag{4}$$

or in column vector format:

$$\boldsymbol{z} := \operatorname{vec}\mathcal{Z} = \sum_{r=1}^{R} \boldsymbol{a}(r) \boxtimes \boldsymbol{b}(r) \boxtimes \boldsymbol{s}(r) \tag{5}$$

3 Parameter identification

3.1 Model identification

It is always possible to decompose the data tensor into a sum of *decomposable* tensors [11, 4] of the form $\mathcal{D}(r) = \boldsymbol{a}(r)\otimes\boldsymbol{b}(r)\otimes\boldsymbol{s}(r)$, that is, in terms of array of coordinates:

$$D_{klm}(r) = a_k(r)b_l(r)s_m(r)$$

Hence tensor $\mathcal{Z}$ takes the form:

$$\mathcal{Z} = \sum_{r=1}^{R} \varsigma_r\, \mathcal{D}(r) \tag{6}$$

where coefficients ς_r can always be chosen to be real positive, and decomposable tensors $\mathcal{D}(r)$ to have unit norm, *i.e.* for L^p norms, $\|\mathcal{D}\| = \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|\,\|\boldsymbol{c}\| = 1$. The minimal value of R such that this decomposition holds is called rank of $\mathcal{Z}$. If R is not too large, the corresponding decomposition is unique [4, 12, 11, 13] and deserves to be referred to as Canonical Polyadic (CP); other terminologies include *rank decomposition* or *Candecomp/Parafac*. Note that decomposable tensors have a rank equal to 1. Because of the uniqueness of the CP decomposition, decomposable tensors of (4) and (6) coincide in the absence of noise. This means that vectors $\{\boldsymbol{a}(r),\ \boldsymbol{b}(r),\ \boldsymbol{s}(r)\}$ coincide up to some scaling factors [4, 8, 11].

3.2 Model identifiability

There exist sufficient conditions ensuring uniqueness of the exact CP, *e.g.* the Kruskal condition [4]:

$$\kappa_A + \kappa_B + \kappa_C \ge 2R + 2 \tag{7}$$

where the notation κ_A refers to the *Kruskal*-rank* of matrix $\boldsymbol{A}$. However, less stringent conditions guaranteeing almost surely a unique solution can be found [12, 11, 13]:

$$R(K + L + M - 2) < KLM$$

*The Kruskal rank of a matrix $\boldsymbol{A}$ is the largest number κ_A such that any subset of κ_A columns are lineraly independent.

This hold true when data are not corrupted by noise. However, if noise is present, we have to solve a *best rank-R approximation problem*:

$$\min_{\boldsymbol{a}_r,\boldsymbol{b}_r,\boldsymbol{s}_r} \left|\left| \mathcal{T} - \sum_{r=1}^{R} \boldsymbol{a}_r \otimes \boldsymbol{b}_r \otimes \boldsymbol{s}_r \right|\right|_F^2 \tag{8}$$

For $d \geq 3$, the best approximation of a d-partite function of a sum of R product of d separable functions does not exist in general [12], as a sequence of rank-r functions can converge to a limit which is not rank-r. A sufficient condition ensuring existence of a solution to (8) via the definition of a physical constraint, the *coherence*, is derived in [12]. Unlike the Kuskal rank, coherences are easy to compute and present the advantage of having a physical meaning, *i.e.* the best rank-R approximation exists and is unique if either impinging signals are not too correlated, or their directions of arrivals and polarization states are not too close.

4 Performances

4.1 Mixed real-complex gradients

Since the parameter of the array processing model are complex, a definition of the derivative of a real function $\boldsymbol{h}(\boldsymbol{z}) \in \mathbb{R}^p$ with respect to the complex variable $\boldsymbol{z} \in \mathbb{C}^n$, $\boldsymbol{z} = \boldsymbol{x} + \imath\, \boldsymbol{y}$, $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}$, needs to be introduced. [7] presents a derivation of Cramér-Rao bounds related to the CP decomposition of multidimensional arrays, using the same definition of complex derivative as in [8, 14]:

$$\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{z}} = \frac{1}{2}\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}} - \frac{\imath}{2}\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{y}} \tag{9}$$

For clarity, our notation of the derivative of a scalar function $\Upsilon(\boldsymbol{z})$ with respect to a column vector $\boldsymbol{z} \in \mathbb{C}^n$ is:

$$\begin{cases} \frac{\partial \Upsilon}{\partial \boldsymbol{z}} \text{is a column vector} \\ \frac{\partial \Upsilon}{\partial \boldsymbol{z}^\mathsf{T}} \text{is a line vector} \end{cases}$$

Given an holomorphic column vector function $\boldsymbol{f}(\boldsymbol{z}) \in \mathbb{C}^m$, we define $[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{z}^\mathsf{T}}]_{ij} = \frac{\partial f_i}{\partial z_j}$ so that

$$\begin{cases} \frac{\partial \boldsymbol{f}^\mathsf{T}}{\partial \boldsymbol{z}} \text{ is an } n \times m \text{ matrix} \\ \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{z}^\mathsf{T}} \text{ is an } m \times n \text{ matrix} \end{cases}$$

In the sequel, we shall need a complex derivative chain rule. Given a scalar function $\Upsilon(\boldsymbol{z}) \in \mathbb{R}$, a complex function $\boldsymbol{z}(\theta) = \boldsymbol{x} + \imath\, \boldsymbol{y} \in \mathbb{C}^p$, and a real variable $\boldsymbol{\theta} \in \mathbb{R}^q$, we have from the real derivative chaine rule:

$$\frac{\partial \Upsilon}{\partial \boldsymbol{\theta}^\mathsf{T}} = \frac{\partial \Upsilon}{\partial \boldsymbol{x}^\mathsf{T}}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}^\mathsf{T}} + \frac{\partial \Upsilon}{\partial \boldsymbol{y}^\mathsf{T}}\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{\theta}^\mathsf{T}}$$

which, using (9), yields the chaine rule:

$$\frac{\partial \Upsilon}{\partial \boldsymbol{\theta}^\mathsf{T}} = 2\Re\left\{ \frac{\partial \Upsilon}{\partial \boldsymbol{z}^\mathsf{T}}\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{\theta}^\mathsf{T}} \right\} \tag{10}$$

4.2 Gradient calculation

In order to compute Cramér-Rao bounds, we shall need the gradients of the log-likelihood, which turn out to be the same as those of the cost function f defined below, deduced from (5), if noise is i.i.d. circular Gaussian:

$$f(\boldsymbol{\vartheta}) = \frac{1}{\sigma_n^2}\left\|\boldsymbol{t} - \boldsymbol{z}(\boldsymbol{\vartheta})\right\|_2^2,\ \boldsymbol{\vartheta} := [\boldsymbol{\theta}; \boldsymbol{\psi}; \boldsymbol{\alpha}; \boldsymbol{\beta}; \operatorname{vec} \boldsymbol{S}; \operatorname{vec} \boldsymbol{S}^*] \tag{11}$$

where σ_n^2 denotes its variance. The gradient expressions will also be subsequently useful to implement descent algorithms. According to the chain rule (10) and definition (9), the partial derivatives of the cost function with respect to DoA parameters are given by

$$\frac{\partial f}{\partial \theta_r} = 2\Re\left\{ \frac{\partial f}{\partial \boldsymbol{a}_r^\mathsf{T}}\frac{\partial \boldsymbol{a}_r}{\partial \theta_r} + \frac{\partial f}{\partial \boldsymbol{b}_r^\mathsf{T}}\frac{\partial \boldsymbol{b}_r}{\partial \theta_r} \right\},\ \frac{\partial f}{\partial \alpha_r} = 2\Re\left\{ \frac{\partial f}{\partial \boldsymbol{b}_r^\mathsf{T}}\frac{\partial \boldsymbol{b}_r}{\partial \alpha_r} \right\}$$
$$\frac{\partial f}{\partial \psi_r} = 2\Re\left\{ \frac{\partial f}{\partial \boldsymbol{a}_r^\mathsf{T}}\frac{\partial \boldsymbol{a}_r}{\partial \psi_r} + \frac{\partial f}{\partial \boldsymbol{b}_r^\mathsf{T}}\frac{\partial \boldsymbol{b}_r}{\partial \psi_r} \right\},\ \frac{\partial f}{\partial \beta_r} = 2\Re\left\{ \frac{\partial f}{\partial \boldsymbol{b}_r^\mathsf{T}}\frac{\partial \boldsymbol{b}_r}{\partial \beta_r} \right\}$$

with

$$\frac{\partial f}{\partial \boldsymbol{a}_r^\mathsf{T}} = \left(\boldsymbol{z} - \sum_{r=1}^{R} \boldsymbol{a}_r \boxtimes \boldsymbol{b}_r \boxtimes \boldsymbol{s}_r \right)^\mathsf{H} (-\boldsymbol{I}_K \boxtimes \boldsymbol{b}_r \boxtimes \boldsymbol{s}_r)$$
$$\frac{\partial f}{\partial \boldsymbol{b}_r^\mathsf{T}} = \left(\boldsymbol{z} - \sum_{r=1}^{R} \boldsymbol{a}_r \boxtimes \boldsymbol{b}_r \boxtimes \boldsymbol{s}_r \right)^\mathsf{H} (-\boldsymbol{a}_r \boxtimes \boldsymbol{I}_L \boxtimes \boldsymbol{s}_r)$$
$$\frac{\partial f}{\partial \boldsymbol{s}_r^\mathsf{T}} = \left(\boldsymbol{z} - \sum_{r=1}^{R} \boldsymbol{a}_r \boxtimes \boldsymbol{b}_r \boxtimes \boldsymbol{s}_r \right)^\mathsf{H} (-\boldsymbol{a}_r \boxtimes \boldsymbol{b}_r \boxtimes \boldsymbol{I}_M)$$

$$\frac{\partial \boldsymbol{a}_r}{\partial \theta_r} = \left[\imath \frac{\omega}{v} (-g_k^x \sin\theta_r \cos\psi_r + g_k^y \cos\theta_r \cos\psi_r)\, A_{kr} \right]_{k=1}^{K}$$
$$A_{kr} = \tfrac{1}{K} \exp\left\{ \imath \tfrac{\omega}{v} [g_k^x \cos\theta_r \cos\psi_r + g_k^y \sin\theta_r \cos\psi_r + g_k^z \sin\psi_r] \right\}$$
$$\frac{\partial \boldsymbol{b}_r}{\partial \theta_r} = \frac{\partial \boldsymbol{V}_r}{\partial \theta_r}\boldsymbol{Q}_r \boldsymbol{w}_r, \qquad \frac{\partial \boldsymbol{b}_r}{\partial \psi_r} = \frac{\partial \boldsymbol{V}_r}{\partial \psi_r}\boldsymbol{Q}_r \boldsymbol{w}_r$$
$$\frac{\partial \boldsymbol{V}_r^{TR}}{\partial \theta_r} = \begin{bmatrix} -\cos\theta_r & \sin\theta_r \sin\psi_r \\ -\sin\theta_r & -\cos\theta_r \sin\psi_r \\ 0 & 0 \end{bmatrix}$$
$$\frac{\partial \boldsymbol{a}_r}{\partial \psi_r} = \left[\imath \frac{\omega}{v} (-g_k^x \cos\theta_r \sin\psi_r - g_k^y \sin\theta_r \sin\psi_r + g_k^z \cos\psi_r)\, A_{kr} \right]_{k=1}^{K}$$
$$\frac{\partial \boldsymbol{V}_r^{TR}}{\partial \psi_r} = \begin{bmatrix} 0 & -\cos\theta_r \cos\psi_r \\ 0 & -\sin\theta_r \cos\psi_r \\ 0 & -\sin\psi_r \end{bmatrix}$$
$$\frac{\partial \boldsymbol{b}_r}{\partial \alpha_r} = \boldsymbol{V}_r \frac{d\boldsymbol{Q}_r}{d\alpha_r}\boldsymbol{w}_r, \qquad \frac{\partial \boldsymbol{b}_r}{\partial \beta_r} = \boldsymbol{V}_r \boldsymbol{Q}_r \frac{d\boldsymbol{w}_r}{d\beta_r}$$
$$\frac{d\boldsymbol{Q}_r}{d\alpha_r} = \begin{bmatrix} -\sin\alpha_r & \cos\alpha_r \\ -\cos\alpha_r & -\sin\alpha_r \end{bmatrix}, \quad \frac{d\boldsymbol{w}_r}{d\beta_r} = \begin{bmatrix} -\sin\beta_r \\ \imath\cos\beta_r \end{bmatrix}$$

4.3 Cramér-Rao Bounds

Cramér-Rao Bounds (CRB) represent the lower bound on the variance of any unbiased estimator of a deterministic parameter. Define the Signal-to-Noise ratio (SNR) as [7]:

$$SNR = 10\log_{10}\frac{\|\mathcal{Z}\|_F^2}{KLM\sigma_n^2}$$

where operator $\|\cdot\|_F^2$ indicates Frobenius norm. For a zero-mean, circularly complex Gaussian noise with covariance $\sigma_n^2 \boldsymbol{I}$ the log-likelihood takes the form (11) up to an additive constant. Then, the mixed real-complex Fisher Information Matrix (FIM) can be shown to be given by [7, 8]:

$$\boldsymbol{\Phi}(\boldsymbol{\vartheta}) = \mathbb{E}\left\{ \left(\frac{\partial f(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} \right)^{\mathsf{H}} \left(\frac{\partial f(\boldsymbol{\vartheta})}{\partial \boldsymbol{\vartheta}} \right) \right\}$$

The CRB of any unbiased estimator of a vector parameter $\boldsymbol{\vartheta}$ is is given by the inverse of the FIM. It is useful to separate parameters to be estimated in three vectors: one real, $[\boldsymbol{\theta}; \boldsymbol{\psi}; \boldsymbol{\alpha}; \boldsymbol{\beta}]$, and two complex, $\operatorname{vec} \boldsymbol{S}$ and $\operatorname{vec} \boldsymbol{S}^*$. With this organization, the FIM has 9 blocks [8]:

$$\boldsymbol{\Phi} = \frac{1}{\sigma_n^2} \begin{pmatrix} 2\Re\{\boldsymbol{G}_{11}\} & \boldsymbol{G}_{12} & \boldsymbol{G}_{12}^* \\ \boldsymbol{G}_{12}^{\mathsf{H}} & \boldsymbol{G}_{22} & \boldsymbol{0} \\ \boldsymbol{G}_{12}^{\mathsf{T}} & \boldsymbol{0} & \boldsymbol{G}_{22}^* \end{pmatrix}$$

where $\boldsymbol{G}_{ij} = \left(\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{\vartheta}_i} \right)^{\mathsf{H}} \left(\frac{\partial \boldsymbol{z}}{\partial \boldsymbol{\vartheta}_j} \right)$.

5 Computer experiments

Signals were simulated according to realistic sampling conditions ($1kHz$ sampling frequency, $M = 42$ time samples, $K = 9$ sensors, $L = 3$ polarization components). Ultimate performances have been evaluated by running a gradient descent initialized with the true values slightly corrupted by noise. A comparison with deterministic CRB is shown in Figure 1. The performance criterion is the *total mean square error* (total MSE) of each DoA and polarization parameter ϑ: $\frac{1}{N}\sum_{n=1}^{N}\sum_{r=1}^{R}(\hat{\vartheta}_{rn} - \vartheta_r)^2$, where $\hat{\vartheta}_{rn}$ is the estimated parameter of the r-th source at the n-th Monte-Carlo trial, $N = 99$ is the number of trials. The number of simultaneous sources was chosen to be $R = 2$, with the following parameters:

$$\begin{cases} \theta_1 = -\pi/3,\ \theta_2 = \pi/6,\ \psi_1 = -\pi/4,\ \psi_2 = \pi/7 \\ \alpha_1 = \pi/5,\ \alpha_2 = \pi/7,\ \beta_1 = -\pi/7,\ \beta_2 = \pi/5 \end{cases}$$

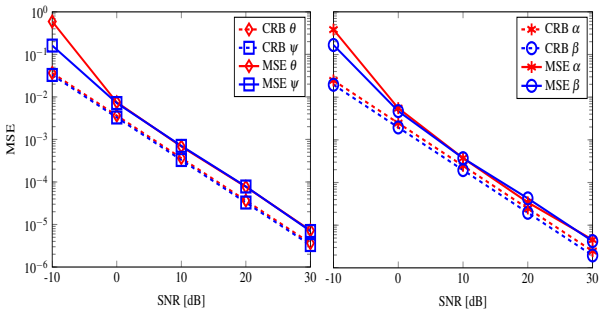

FIG. 1: MSE vs SNR

CRBs are obtained by summing the diagonal entries in the inverse of the first block, $2\Re\{\boldsymbol{G}_{11}\}$, in the FIM. This means source signals were considered as known and not as nuisances (which implies the obtention of slightly smaller bounds).

References

[1] S. Anderson and A. Nehorai, "Analysis of a polarized seismic wave model," *IEEE Trans. Sig. Proc.*, vol. 44, no. 2, pp. 379–386, 1996.

[2] A. Nehorai and E. Paldi, "Vector-sensor array processing for electromagnetic source localization," *IEEE Trans. Sig. Proc.*, vol. 42, no. 2, pp. 376–398, 1994.

[3] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Trans. Sig. Proc.*, vol. 42, no. 9, pp. 2481–2491, 1994.

[4] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Sig. Proc.*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.

[5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antenna Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[6] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech Sig. Proc.*, vol. 37, no. 7, pp. 984–995, 1989.

[7] X. Liu and N. D. Sidiropoulos, "Cramér-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Trans. Sig. Proc,*, vol. 49, no. 9, pp. 2074–2086, 2001.

[8] S. Sahnoun and P. Comon, "Tensor polyadic decomposition for antenna array processing," in *21st Int. Conf. Comput. Stat (CompStat)*, Geneva, Aug. 19-22 2014, pp. 233–240, hal-00986973.

[9] X. Guo, S. Miron, D. Brie, S. Zhu, and X. Liao, "A Candecomp/Parafac perspective on uniqueness of doa estimation using a vector sensor array," *IEEE Trans. Sig. Proc.*, vol. 59, no. 7, pp. 3475–3481, 2011.

[10] X. Guo, S. Miron, and D. Brie, "The effect of polarization separation on the performance of Candecomp/Parafac-based vector sensor array processing," *Physical Communication*, vol. 5, no. 4, pp. 289–295, 2012.

[11] P. Comon, "Tensors: a brief introduction," in *IEEE Sig. Proc. Magazine*, 2014, vol. 31, pp. 44–53.

[12] L.-H. Lim and P. Comon, "Blind multilinear identification," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1260–1280, Feb. 2014, open access.

[13] L. Chiantini, G. Ottaviani, and N. Vannieuwenhoven, "An algorithm for generic and low-rank specific identifiability of complex tensors," *SIAM J. Matrix Ana. Appl.*, vol. 35, no. 4, pp. 1265–1287, 2014.

[14] A. Hjorungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Sig. Proc.*, vol. 55, no. 6, pp. 2740–2746, 2007.